Your thesaurus codes are:
(12.12.1; Universe 11.03.1;

ASTROPHYSICS
December 4, 2000

Detection of non-random patterns in cosmological gravitational clustering

R. Valdarnini[1]

SISSA, via Beirut 2-4, 34014, Trieste, Italy

Received ...; accepted ...

Abstract. A new method for analyzing point patterns produced by the evolution of gravitational clustering is presented. The method is taken from the study of molecular liquids, where it has been introduced for making a statistical description of anisotropic distributions. The statistical approach is based on the spherical harmonic expansion of angular correlations. A general introduction to the method is given; a theoretical analysis shows that it is partially connected with previous harmonic analyses applied to galaxy catalogs. The effectiveness of the statistical analysis in quantifying clustering morphology is illustrated by applying the statistical estimators to point distributions produced by an ensemble of cosmological $N-$ body simulations with a CDM spectrum. The results demonstrate that the statistical method is able to detect anisotropies in non-random patterns, with different scales being probed according to the expansion coefficients.

Key words: large-scale structure of Universe–clustering

1. Introduction

According to the standard picture the observed structures in the Universe have arisen via gravitational instability from the time evolution of the initial matter density fluctuation field $\delta_m(\boldsymbol{x})$. At very early epochs the field is assumed to be a random Gaussian process (e.g. Peebles 1980). In the proposed scenarios the formation of structures proceeds hierarchically, with galaxies forming first and larger structures following later from clustering of galaxies. The present galaxy distribution shows that clustering is ranging from small groups up to clusters and superclusters. A striking feature of the observed large-scale galaxy distribution, as revealed by extended ($\simeq 100-200Mpc$) redshift surveys (Geller & Huchra 1989; da Costa et al. 1994; Landy et al. 1996; Broadhurst et al. 1990; Saunders et al. 1991a), is that galaxies are connected in a web of sheet-like structures, with voids between them and filaments at the intersections (Geller & Huchra 1989). The size of these structures can be as large as $50 \simeq 100Mpc$. These features revealed in the galaxy distribution have been formed through gravitational clustering and are connected to the initial power spectrum $P(k) =< |\delta_m(k)|^2 >$; thus the observed galaxy clustering can in principle be used to put constraints on the allowed cosmological models, for which the fundamental parameters determine the shape of $P(k)$. The galaxy distribution today is highly non-linear and $N-$body methods have been used to sample the initial phase space distribution and to study the time evolution of the clustering growth (Efstathiou et al. 1985; Bertschinger & Gelb 1991). These methods lead to a final particle distribution which should be a representative sample of the expected galaxy distribution for the cosmological model under study. The use of $N-$body codes implicitly makes the assumption that galaxies trace the matter distribution. In order to compare the observed galaxy clustering with the results of numerical simulations one has to introduce a statistical descriptor. The method used to investigate the statistical distribution of galaxies is therefore very important because it should be used to discriminate between different models, furthermore the method must also provide a mathematical description for the rich variety of clustering morphology as seen in the galaxy distribution.

The first method to be introduced as a tool for studying galaxy clustering was the 2-point correlation function $\xi(r)$ (Totsuji & Kihara 1969; Peebles & Hauser 1974). Analyses of galaxy catalogs show that $\xi(r)$ has a power-law shape $\xi \propto r^{-\gamma}$, with $\gamma = 1.77$ and is equal to unity for $r_o = (5 \pm 1)h^{-1}Mpc$ [1] (Peebles & Hauser 1974; Fisher at al. 1994). The 2-point correlation function has been widely used to constrain cosmological models from results of numerical simulations (Jenkins et al. 1998 and references cited therein). However the 2-point function does not provide useful information about the rich variety of structures which characterize the morphology of the galaxy distributions. The reason is that the power spectrum $P(k)$ (the Fourier transform of ξ) does not describe the correlation of the phases that the $\delta_{\mathbf{k}}$'s develop during the clustering evolution.

[1] Here $H_0 = 100hKmsec^{-1}Mpc^{-1}$ is the present value of the Hubble constant

For a complete description of the clustering one has in principle to consider correlation functions ξ_N of higher order. The measurement of these correlations is impractical however when $N \geq 5$ because of the noise introduced by the finite number of galaxies in the sample (Szapudi & Colombi 1996). For this reason different statistical tools have been introduced for studying the clustering of the galaxy distribution. These methods are : the void probability distribution $P_0(V)$ (White 1979; Vogeley et al. 1994; Ghigna et al. 1997); moments of counts in cells (Saunders et al. 1991b); probability distribution of count in cells (Bouchet et al. 1993; Ueda & Yokoyama 1996).

While these methods still make use of the higher order moments of the clustering distribution, alternative approaches have been introduced which are more geometrical and implicitly contain information from the $N-$ point correlations at all orders. The first method to be introduced was the percolation statistic (Zel'dovich 1982; Klypin & Shandarin 1993), other methods include the topological genus-density threshold (Gott, Dickinson & Melott 1986; Ryden et al. 1989), minimal spanning trees (Barrow, Bhavsar & Sonoda 1985); Minkowski functionals (Mecke, Buchert & Wagner 1994; Kerscher et al. 1997), graph-theory (Ueda & Itoh 1999) or a global descriptor (J-function) based on the nearest-neighbor distribution (Kerscher at al. 1999). In order to quantify the geometrical features of the galaxy clustering several statistics have been explicitly developed to single out in a quantitative way the filaments and sheets of the cosmic network. These statistics are based on the moments of the mass distribution and can be used to detect filaments (Fry 1986; Davé et al. 1999; Babul & Starkman 1992), or the shape of the clustering (Luo & Vishniac 1995; Robinson & Albrecht 1996). The shape statistic can also be constructed using Minkowski functionals (Schmalzing, & Buchert 1997; Bharadwaj et al. 2000).

The existence of this variety of approaches is related to the fact that there is not any single well defined theoretical method which can be used to analyze the shape of the clustering network. Furthermore a statistical method must also be able to produce a robust statistical measure, in order to be applied to galaxy catalogs. These requirements are naturally satisfied by the 2-point function ξ, but, as outlined above, this function it is not indicated for detecting filaments.

In this paper, I propose an alternative method for analyzing the clustering morphology which is a generalization of the 2-point function ξ and is based on the spherical harmonic analysis. The statistical method has been applied in molecular fluid dynamic simulations for describing the structures of anisotropic fluids or disordered metallic glasses (Steinhardt, Nelson & Ronchetti 1983; Wang & Stroud 1991). The method is applied here for illustrative purposes to point distributions obtained from cosmological $N-$body simulations.

The paper is organized as follows: In Sect.2, I present the method. Sect.3 describes the ensemble of CDM cosmological simulations used to test the effectiveness of the method. In Sect.4 the statistic is applied to the point sets produced by the simulations, the main results are discussed and the conclusions are summarized.

2. THE METHOD

2.1. Spherical harmonic analysis

For describing galaxy clustering evolution in the nonlinear regime, use is made of different statistical methods, based on the galaxy positions, which have been outlined in the Introduction. The approach introduced here makes use of the positions of the point distribution within a certain distance from a randomly chosen one.

The method is drawn from molecular dynamics simulations, where it has been introduced for studying orientational order of supercooled liquids and metallic glasses. A general review can be found in Haile & Gray (1980) and Mc Donald (1986), here I will follow the notation of Wang & Stroud (1991).

Let us consider a system of N_p particles. The $i-th$ particle has coordinates $\boldsymbol{r}_i$, in an arbitrary reference frame. For a specified cutoff radius R_c all the particles such that $|\boldsymbol{r}_i - \boldsymbol{r}_j| < R_c$ are neighbors of i.

The line joining i to one of the j is termed a *bond*.

The angular coordinates of the vector $\boldsymbol{\Delta}_{ji} \equiv \boldsymbol{r}_j - \boldsymbol{r}_i$ are θ_j, ϕ_j and the quantity

$$Q_{lm}(\boldsymbol{r}_i) = \sum_{j \neq i} Y_{lm}(\theta_j, \phi_j), \tag{1}$$

is the coefficient of the spherical harmonic expansion of the angular density of the bonds associated with the particle i.

In Eq. 1, and hereafter, summation is understood over all particles j of the distribution such that $|\boldsymbol{r}_i - \boldsymbol{r}_j| < R_c$.

The coefficients $Q_{lm}(\boldsymbol{r}_i)$ are defined as the bond-orientational order parameters (Wang & Stroud 1991) and they can be drastically changed by a rotation of the reference systems. A natural quantity to consider , which is rotation invariant, is

$$Q_l(\boldsymbol{r}_i) = \sqrt{\frac{4\pi}{2l+1} \sum_{m=-l}^{m=l} Q^{\star}_{lm}(\boldsymbol{r}_i) Q_{lm}(\boldsymbol{r}_i)}. \tag{2}$$

Using the addition theorem for the spherical harmonics, the expression simplifies to

$$Q_l(\boldsymbol{r}_i) = \sqrt{\sum_j \sum_k P_l(\gamma_{jk})}, \tag{3}$$

where P_l is the Legendre polynomial,

$$\gamma_{jk} \equiv cos(\theta_{jk}) = \mathbf{\Delta}_{ji} \cdot \mathbf{\Delta}_{ki}/(|\mathbf{\Delta}_{ji}||\mathbf{\Delta}_{ki}|)$$

is the angle between two bonds and the summations in Eq. 3 are then independent of the chosen frame.

The expression for $Q_l(\boldsymbol{r}_i)$ can be further simplified if one considers that $\gamma_{jj} = 1$, furthermore it is more convenient to redefine $Q_l(\boldsymbol{r}_i)$ dividing by the mean number of neighbors $< N_i >= \bar{N}_i$ of the particle i :

$$Q_l(\boldsymbol{r}_i) = \frac{1}{\bar{N}_i}\sqrt{N_i + 2\sum_j \sum_{k>j} P_l(\gamma_{jk})}. \quad (4)$$

For a random distribution the summation terms in the square root are negligible for $N_i \gg 1$, thus $Q_l(\boldsymbol{r}_i) \propto 1/\sqrt{N_i}$ for Poisson noise.

For the whole system an order parameter $< Q_l >$ can be defined by averaging over all of the N_p particles:

$$< Q_l >= \frac{1}{N_p}\sum_{i=1}^{N_p} Q_l(\boldsymbol{r}_i) \equiv Q_l. \quad (5)$$

Thus the clustering distribution can be studied by evaluating the Q_l at a specified cutoff radius. Going to higher coefficients ($l \gg 1$) means probing smaller scales in the bond distribution.

However the Q_l do not exhibit spatial information, and a more useful quantity is the auto-correlation $G_l(r)$ of the coefficients $Q_l(\boldsymbol{r}_i)$. The function G_l is defined as follows: for all of the M_p pairs (i,k), such that $|\boldsymbol{r}_i - \boldsymbol{r}_k| = r \pm \Delta r$, where Δr is the thickness of the radial bin, then $G_l(r)$ is the sum over all of these pairs

$$G_l(r) = \frac{1}{M_p}\sum_i \sum_k \frac{4\pi}{2l+1}\sum_{m=-l}^{m=l} Q^\star_{lm}(\boldsymbol{r}_i)Q_{lm}(\boldsymbol{r}_i + \boldsymbol{r}). \quad (6)$$

This equation can be greatly simplified: let j be the set of neighbors of the particle i and p that of the particle k, which satisfy $|\boldsymbol{r}_j - \boldsymbol{r}_i| < R_c$ and $|\boldsymbol{r}_p - \boldsymbol{r}_k| < R_c$. Then the summation becomes

$$\sum_{m=-l}^{m=l} Q^\star_{lm}(\boldsymbol{r}_i)Q_{lm}(\boldsymbol{r}_i + \boldsymbol{r}) = \sum_{m=-l}^{m=l} Q^\star_{lm}(\boldsymbol{r}_i)Q_{lm}(\boldsymbol{r}_k) = \sum_{m=-l}^{m=l}\sum_j Y^\star_{lm}(\theta_j,\phi_j)\sum_p Y_{lm}(\theta_p,\phi_p).$$

The final expression for G_l is:

$$G_l(r) = \frac{1}{M_p}\sum_i\sum_k\sum_j\sum_p P_l(\Gamma_{jp}), \quad (7)$$

with Γ_{jp} being the angle between $\mathbf{\Delta}_{ji}$ and $\mathbf{\Delta}_{pk}$. The summation over the pairs is $\sum_i \sum_k$, with the sum over the particles k only for those particles with $|\boldsymbol{r}_i - \boldsymbol{r}_k| = r \pm \Delta r$.

Note that now it is not possible to simplify the equation as was done for Eq. 4, because one is considering the correlation between *different* order parameters.

The information about the degree of angular correlation between bonds contained in the $G_l(r)$ is more explicit if the G_l themselves are divided by the zero-order correlation function

$$G_0(r) = \frac{4\pi}{M_p}\sum_i\sum_k Q^\star_0(\boldsymbol{r}_i)Q_0(\boldsymbol{r}_i + \boldsymbol{r}). \quad (8)$$

The functions $G_l(r)$ can be made dependent also on the direction of $\boldsymbol{r}$ by considering all the pairs with orientation of the relative separation $\boldsymbol{r}$ in the solid angle range $\Omega_{\mathsf{r}} \pm \Delta\Omega_{\mathsf{r}}$. These functions $G_l(\boldsymbol{r})$ are of scarce practical use because of the noise level which is introduced when applied to a finite sample.

2.2. Theory

The equations that have been obtained for Q_l and $G_l(r)$ are defined through a point set distribution; from a theoretical point of view it is possible to relate these functions to the power spectrum of the density fluctuations. Spherical harmonic analysis has been widely applied in cosmology, both to the galaxy distribution (Scharf et al. 1992; Scharf et al. 1993; Ballinger, Heavens & Taylor 1995) and to the study of cosmic background radiation maps (Hu, Sugiyama & Silk 1997; Bartlett 1999). I will follow here the notation of Scharf et al. (1992), who have analyzed the angular distribution of the IRAS redshift survey using spherical harmonics. For the generic distribution considered in Sect.2.1 let us define $n(\boldsymbol{r})$ to be the number density and $\bar{n}$ to be the average one. It is also useful to introduce $\Phi(r)$ as a generic selection function, here $\Phi = \bar{n}$.

Then the coefficients $Q^m_l(\boldsymbol{r}_o)$, $\boldsymbol{r}_o$ being the center position, are given by

$$Q^m_l(\boldsymbol{r}_o) = \int \Phi(\boldsymbol{r'})\frac{n(\boldsymbol{r}^{'})}{\bar{n}}Y^m_l(\Omega_{\mathsf{r}'})d^3r^{'} \qquad \boldsymbol{r} = \boldsymbol{r}_o + \boldsymbol{r'},$$

where the integral is taken over a sphere of radius R_c. The number density is

$$n(\boldsymbol{r}) = \bar{n}[1 + \delta(\boldsymbol{r})] = \bar{n}[1 + \delta(\boldsymbol{r}_o + \boldsymbol{r}^{'})],$$

here $\delta(\boldsymbol{r})$ is the density fluctuation. Thus the equation for $Q^m_l(\boldsymbol{r}_o)$ becomes

$$Q^m_l(\boldsymbol{r}_o) = \int \Phi[1+\delta]Y^m_l d^3r^{'} = \int \Phi Y^m_l d^3r' + \int \Phi(\boldsymbol{r}^{'})\delta(\boldsymbol{r}_o + \boldsymbol{r}^{'})Y^m_l d^3r^{'},$$

where the first term on the rhs is zero for $l \geq 1$. The density fluctuation $\delta(\boldsymbol{r})$ can be expanded into Fourier modes, and the final expression for $Q^m_l(\boldsymbol{r}_o)$ is

$$Q^m_l(\boldsymbol{r}_o) = \sum_{\mathsf{k}} \delta_{\mathsf{k}}\, e^{i\mathsf{k}\mathsf{r}_o}\int \Phi(\boldsymbol{r}^{'})e^{i\mathsf{k}\cdot r^{'}}Y^m_l(\Omega_{\mathsf{r}'})d^3r^{'}. \quad (9)$$

The $e^{i\mathsf{k}\mathsf{r}}$ is now decomposed into spherical functions, so that the ensemble average gives

$$< |Q_l^m|^2 > \; = \; (4\pi)^2 \sum_{\mathsf{k}} |\delta_k|^2 \left(\int_0^{R_c} \Phi(x) x^2 j_l(kx) dx \right)^2 |Y_l^m(\Omega_{\mathsf{k}})|^2, \quad (10)$$

here the $j_l(x)$s are the spherical Bessel functions. The summation over k can be translated into an integral and $\int |Y_l^m|^2 d\Omega = 1$, so that

$$< |Q_l^m|^2 > \; = \; V \frac{2}{\pi} \int_0^\infty P(k) k^2 dk \left[\int_0^{R_c} \Phi(x) x^2 j_l(kx) dx \right]^2 + \mathcal{N}, \quad (11)$$

where $P(k) =< |\delta_k|^2 >$, V is a normalizing volume and $\mathcal{N}$ is the shot noise term $\mathcal{N} = \int \Phi(x) x^2 dx$.

The coefficients C_l are defined as

$$C_l^2 = \frac{1}{\mathcal{N}(\Omega/4\pi)(2l+1)} \sum_m |Q_l^m|^2, \quad (12)$$

and assuming statistical isotropy they can be related to Eq.(5) by $Q_l \simeq C_l / \bar{N}_i^{1/2}$. For an uncorrelated distribution $C_l = 1$, furthermore the integral $\int_0^{R_c} \Phi(x) x^2 j_l(kx) dx \equiv \Psi_l(kR_c)$ defines a bidimensional window function in the parameter space (R_c, l). Increasing the value of l means probing smaller angular scales, the value of the cutoff radius R_c sets the size of the spectrum contribution to the integral (see, for example, Fig. 4 of Scharf et al. 1993). The statistical method introduced in Sect.IIA, and originally proposed in a different field, is then mathematically equivalent to the spherical harmonic analysis which has been applied to galaxy surveys (Scharf et al. 1993). The order parameter Q_l is just related to the angular power spectrum estimator C_l (Scharf et al. 1992).

However Eq. (10) is valid as long as the position of the observer located at $\boldsymbol{r}_0$ can be considered as random with respect to the bond distribution inside the sphere of radius R_c. This is a good approximation for values of the cut-off radius R_c much larger than the 2-point clustering length, a condition which is valid for the analyzed angular catalogs. In Sect. 4 the statistical method is tested using the point distributions obtained from a set of cosmological simulations, and the coefficients Q_l are calculated for different parameters and cut-off radii R_c. Therefore one has to consider in principle also the three-point correlation terms between the particle located at $\boldsymbol{r}_0$ and the other particles in the bond sphere. The ensemble average for $|Q_l^m|^2$ is then (Peebles 1980, sect. 36B):

$$< |Q_l^m|^2 > \; = \; \int \Phi d\boldsymbol{x} \int \Phi d\boldsymbol{y} [1 + \xi(x) + \xi(y) + \xi(|\boldsymbol{x} - \boldsymbol{y}|) + \zeta(x, y, |\boldsymbol{x} - \boldsymbol{y}|)] Y_l^m(\Omega_{\mathsf{x}}) Y_l^{\star m}(\Omega_{\mathsf{y}}), \quad (13)$$

where $\boldsymbol{x}$ and $\boldsymbol{y}$ are the particle separations relative to the origin and ζ is the reduced three-point correlation function. Integration over solid angles removes for $l \geq 1$ the first three terms on the rhs of the equation. Of the remaining terms, the first ($\xi(|\boldsymbol{x} - \boldsymbol{y}| = \sum_{\mathsf{k}} |\delta_k|^2 e^{i\mathsf{k}(\mathsf{x}-\mathsf{y})}$), gives Eq. (10), while the other takes into account for the triangle configuration the correlations between the three particles. The integral can be evaluated in terms of the Fourier transform of ζ (Fry 1984), the bispectrum $B(\boldsymbol{k}_1, \boldsymbol{k}_2, \boldsymbol{k}_3)$:

$$\int \Phi d\boldsymbol{x} \int \Phi d\boldsymbol{y} \zeta(x, y, |\boldsymbol{x} - \boldsymbol{y}|) Y_l^m(\Omega_{\mathsf{x}}) Y_l^{\star m}(\Omega_{\mathsf{y}}) =$$

$$\sum_{\mathsf{k}_1 \mathsf{k}_2 \mathsf{k}_3} B(\boldsymbol{k}_1, \boldsymbol{k}_2, \boldsymbol{k}_3) \int \Phi d\boldsymbol{x} \int \Phi d\boldsymbol{y} Y_l^m(\Omega_{\mathsf{x}}) Y_l^{\star m}(\Omega_{\mathsf{y}})$$

$$e^{i(\mathsf{k}_1\mathsf{x}_1 + \mathsf{k}_2\mathsf{x}_2 + \mathsf{k}_3\mathsf{x}_3)} = \sum_{\mathsf{k}_2 \mathsf{k}_3} B(\boldsymbol{k}_1 = -(\boldsymbol{k}_2 + \boldsymbol{k}_3), \boldsymbol{k}_2, \boldsymbol{k}_3)$$

$$\int \Phi d\boldsymbol{x} \int \Phi d\boldsymbol{y} Y_l^m(\Omega_{\mathsf{x}}) Y_l^{\star m}(\Omega_{\mathsf{y}}) e^{i(\mathsf{k}_2\mathsf{x} + \mathsf{k}_3\mathsf{y})} =$$

$$(4\pi)^2 (-)^l \sum_{\mathsf{k}_2 \mathsf{k}_3} \int \int B(\boldsymbol{k}_1, \boldsymbol{k}_2, \boldsymbol{k}_3) \Psi_l(k_1 R_c) \Psi_l(k_2 R_c) Y_l^m(\Omega_{\mathsf{k}_1}) Y_l^{\star m}(\Omega_{\mathsf{k}_2}) \; , \quad (14)$$

where the constraint $\sum \boldsymbol{k}_i = 0$ is required by homogeneity, $\boldsymbol{x} \equiv \boldsymbol{x}_2 - \boldsymbol{x}_1$ and $\boldsymbol{y} \equiv \boldsymbol{x}_3 - \boldsymbol{x}_1$. In second order perturbation theory, for Gaussian initial conditions, the bispectrum can be expressed as (Fry 1984 ; Verde, Heavens & Matarrese 2000) :

$$B(\boldsymbol{k}_1, \boldsymbol{k}_2, \boldsymbol{k}_3) = K(\boldsymbol{k}_1, \boldsymbol{k}_2) P(k1) P(k2) + cyc. \; , \quad (15)$$

here *cyc.* means cyclic terms. The term $K(\boldsymbol{k}_1, \boldsymbol{k}_2)$ is weakly dependent on the cosmology and can be written as (Fry 1984; Goroff et al. 1986; Verde, Heavens & Matarrese 2000):

$$K(\boldsymbol{k}_1, \boldsymbol{k}_2) = A_0 + A_1 cos(\theta_{12}) + A_2 cos^2(\theta_{12}), \quad (16)$$

where the coefficients A_i depend on the assumed biasing expansion for the local density field and θ_{12} is the angle between $\boldsymbol{k}_1$ and $\boldsymbol{k}_2$. The integral (14) can then be evaluated for a particular cosmological model. The dependence of B on the triangle shape in the k-space shows that the integral is non zero only for $l \leq 2$. The integral (14) is a particular case of the integrals considered by Verde, Heavens and Matarrese (2000). These authors have expanded the projected galaxy density in spherical harmonics and studied the 3-point function of the expansion coefficients, which is a quantity directly related to the bispectrum. The evaluation of the above integral is rather complicated (see, e.g., Eq. (28) of Verde, Heavens and Matarrese 2000) and will not be considered here. In Sect. 4 it will be seen that the corrections (14) to Eq. (10) must be rather small already for $R_c \simeq 2r_0$.

This theoretical analysis can also be used to derive an expression for the quantity $G_l(\boldsymbol{r})$, which contains spatial information about the bond angular distribution. The

evaluation of $G_l(\boldsymbol{r})$ now involves correlation terms up to the fourth order. In analogy with Eq. (13) one has to compute for $G_l(\boldsymbol{r})$ the ensemble average

$$G_l(\boldsymbol{r}) =< Q_l^m(\boldsymbol{r}_o)Q_l^{\star m}(\boldsymbol{r}_o+\boldsymbol{r}) >= \int \Phi d\boldsymbol{x} \int \Phi d\boldsymbol{y}[\xi(|\boldsymbol{r}+\boldsymbol{y}-\boldsymbol{x}|)+\zeta(x,|\boldsymbol{r}+\boldsymbol{y}-\boldsymbol{x}|,|\boldsymbol{r}+\boldsymbol{y}|) +\zeta(|\boldsymbol{r}+\boldsymbol{y}-\boldsymbol{x}|,y,|\boldsymbol{r}-\boldsymbol{x}|)+\xi(|\boldsymbol{r}+\boldsymbol{y}|)\xi(|\boldsymbol{x}-\boldsymbol{r}|)+\xi(r) \xi(|\boldsymbol{r}+\boldsymbol{y}-\boldsymbol{x})|)+\eta(x,r,y,|\boldsymbol{r}+\boldsymbol{y}|,|\boldsymbol{r}-\boldsymbol{x}|,|\boldsymbol{r}+\boldsymbol{y}-\boldsymbol{x}|)] Y_l^m(\Omega_{\mathsf{x}})Y_l^{\star m}(\Omega_{\mathsf{y}}), \quad (17)$$

where η is the reduced four-point correlation function, and only the terms which are non-zero after the integration over the solid angles are shown in the integrals; the separations of the four-particle configuration of Eq. (7) are defined as: $\boldsymbol{r}=\boldsymbol{r}_k-\boldsymbol{r}_i,\ \boldsymbol{y}=\boldsymbol{r}_p-\boldsymbol{r}_k,\ \boldsymbol{x}=\boldsymbol{r}_j-\boldsymbol{r}_i$. The integrals in (17) can be evaluated using the same method adopted for the integral (14). As previously discussed, from the results of Sect. 4 for $< |Q_l^m|^2 >$, the third and fourth-order terms are expected to be negligible. I will consider here only the leading order term $\xi(|\boldsymbol{r}+\boldsymbol{y}-\boldsymbol{x}|)$ and the ensemble average for $G_l(\boldsymbol{r})$ is then

$$< Q_l^m(\boldsymbol{r}_o)Q_l^{\star m}(\boldsymbol{r}_o+\boldsymbol{r}) >= \sum_{\mathsf{k}} e^{-i\mathsf{k}\mathsf{r}}|Y_l^m(\Omega_k)|^2\left[\int \Phi j_l(kx)x^2dx\right]^2(4\pi)^2|\delta_k|^2. \quad (18)$$

The summation over $\boldsymbol{k}$ becomes an integral and if $e^{-i\mathsf{k}\mathsf{r}}$ is now decomposed as

$$e^{-i\mathsf{k}\mathsf{r}} = 4\pi\sum_{pq} i^p j_p(kr)Y_{pq}^*(\Omega_{\mathsf{r}})Y_{pq}(\Omega_{\mathsf{k}}), \quad (19)$$

then the triple integral for the spherical harmonics Y_l^m in Eq.(18) is

$$\int Y_{pq}Y_{lm}^*Y_{lm}d\Omega_k = \int Y_{pq}(-)^m Y_{l-m}Y_{lm}d\Omega_k = (-)^m\left[\frac{(2l+1)^2\ (2p+1)}{4\pi}\right]^{1/2}\begin{pmatrix} l & l & p \\ 0 & 0 & 0\end{pmatrix}\begin{pmatrix} l & l & p \\ m & -m & q\end{pmatrix}.$$

The symbols $\begin{pmatrix} l_1 & l_2 & l_3 \\ m_1 & m_2 & m_3\end{pmatrix}$ are the Wigner $3-j$ coefficients. Because of their properties, the summation over the q terms in Eq. (18) is non-zero only for $q=m-m=0$, furthermore the p summation admits only terms with $2l+p=even, 0\le p\le 2l$. Finally the expression for $G_l(\boldsymbol{r})$ becomes :

$$G_l(\boldsymbol{r}) = \frac{4\pi}{2l+1}\sum_m < Q_l^m(\boldsymbol{r}_o)Q_l^{\star m}(\boldsymbol{r}_o+\boldsymbol{r}) >= \frac{(4\pi)^4}{(2\pi)^3}V\int_0^\infty dk k^2\left[\int_0^{R_c}\Phi x^2 j_l(kx)dx\right]^2 P(k) \sum_{p(even)=0}^{2l} i^p j_p(kr)Y_{p0}^\star(\Omega_{\mathsf{r}})\sum_{m=-l}^{m=l}(-)^m\frac{(2p+1)^{1/2}}{(4\pi)^{1/2}} \begin{pmatrix} l & l & p \\ 0 & 0 & 0\end{pmatrix}\begin{pmatrix} l & l & p \\ m & -m & 0\end{pmatrix}. \quad (20)$$

In the above equation the function G_l still depends on the orientation of $\boldsymbol{r}$ through the m terms. However, owing to statistical isotropy, this dependence should not be present. In fact, it can be shown that (Gangui 1995) $\sum_m(-)^m\begin{pmatrix} l & l & p \\ m & -m & 0\end{pmatrix} = (2l+1)^{(1/2)}(-)^l\delta_{p,0}$, and Eq. (20) simplifies to

$$G_l(r) = \frac{(4\pi)^3}{(2\pi)^3}V\int_0^\infty dkP(k)k^2 \left[\int_0^{R_c}\Phi(x)x^2j_l(kx)dx\right]^2 j_0(kr). \quad (21)$$

From this function one can recover the standard definition for $\xi(r)$ taking for $l=0$ the limit $R_c\to 0, \int_0^{R_c}\Phi(x)x^2j_0(kx)dx\to 1/4\pi$. The statistical descriptor of clustering defined by Eqs. (5) and (7), when applied to a spatial point process, should be able to describe in a quantitative way the patterns generated by the clustering distribution. As outlined in the Introduction different methods have been employed to this end; the most important advantage of the order parameters introduced here is that they can can be used to study spatial patterns in a well-defined way by varying the angular scale l or the depth R_c of the bond spheres. In order to test the clustering features described by the order parameter Q_l and correlation function G_l , I have then analyzed point distributions generated by a set of cosmological N-body simulations with CDM spectra. The simulations are evolved in time starting from an early epoch when linear theory applies. The clustering evolution gives rise to a variety of spatial structures and the statistical descriptors are then applied, in order to analyze the produced patterns, to numerical outputs selected at different epochs. The results obtained are discussed in Section 4.

3. THE SIMULATIONS

The point distributions used for the statistical analysis have been constructed from a set of purely gravitational cosmological N-body simulations. For each of these simulations the clustering evolution of $N_p=10^6$ particles is followed in time in a box of comoving size $L=200h^{-1}Mpc$. The gravitational potential is solved using a P^3M code (Efstathiou et al. 1985), with $N_m=256^3$ grid points. The interparticle force has a linear shape cloud profile with a comoving softening parameter $\varepsilon=0.220Mpc$.

The initial conditions for the particle positions and velocities have been generated at the initial time t_i according to the standard Zel'dovich approximation (Efstathiou et al. 1985). The power spectrum $P(k)$ is that of a standard CDM model (Bardeen et al. 1986) with $\Omega_m=1$ and $h=0.5$. The Fourier components δ_k are Gaussian distributed, with random phases and variance $P(k)$. An ensemble of five different integrations has been

carried out, each with a different random realization for δ_k. The simulations are integrated in time with the expansion factor $a(t)$ being used as the integration variable and with $a(t_i) = 1$. Records of particle positions have been taken at $a(t) = 3.0, 3.8, 4.5, 6, 7.9$. The power spectrum is normalized so that the rms mass fluctuation in a $R = 8h^{-1}Mpc$ sphere takes the value $\sigma_8 = 0.4$ at $a(t) = 3.0$.

In order to measure the order parameters associated with peaks above a given height of the underlying density field $\delta(\boldsymbol{x})$ I have used the peak-background formalism (White et al. 1987). For each particle i the peak number n_i is calculated at $a(t_i)$ from the initial density field; n_i is the number of peaks with height $\delta_s \geq \nu\sigma_s$, here δ_s is the field $\delta(\boldsymbol{x})$ Gaussian smoothed with $R_s = 1.1Mpc$ and σ_s^2 is the variance of δ_s. The field is constrained to take the value $\nu_b\delta_b$ when smoothed on a scale $R_b > R_s$ (White et al. 1987; Park 1991) The particle i is then identified as a 'peak' if $n_i > p$, with p being a uniform random number between 0 and 1.

The total number of particles associated with the peaks of a given height is $\simeq \sum n_i$. Three different threshold levels have been chosen: $\nu = 0.5, 1.3, 2.5$ which correspond, respectively, to $\sum n_i \simeq 250,000,\ 165,000,\ 33,000$ in the simulation box. Hereafter the notation $\{x\}_s$ means the subset of the 10^6 simulation particles selected to identify the peaks with threshold $\nu = s$. At different epochs $a(t)$ the order parameters have been computed by applying the statistical estimators to the point distribution produced by $\{x\}_s$. The clustering of the mass distribution produced by all of the simulation particles has been also considered, in this case $\delta \geq s = -1$.

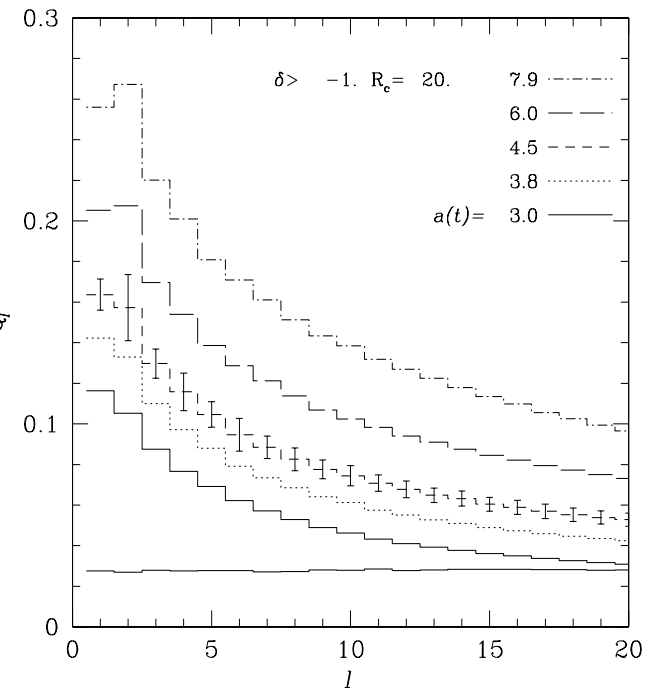

Fig. 1. The coefficients Q_l are plotted as functions of l for different expansion factors $a(t)$. The estimator (5) has been applied to the whole point distribution of the simulation particles. Here and in the following plots the cut-off radius R_c is in units of $h^{-1}Mpc$. Error bars show the scatter within the five simulation ensemble. The continuous line with a nearly constant value for the Q_l refers to a random test distribution.

4. RESULTS AND CONCLUSIONS

The order parameters Q_l are shown in Figs. 1, 2 & 3 for different values of the expansion factor $a(t)$, threshold levels and cut-off radius R_c. In Fig. 1 the time evolution of Q_l is plotted as a function of l for different epochs. Several generic features can be seen in the plot: the coefficients Q_l grow with time as the clustering develops and decay with increasing l as smaller angular scales are probed. This follows directly form the shape of $P(k)$. For a random test distribution with $N_i \simeq 10^3$ neighbors $Q_l \simeq const \simeq 1/\sqrt{N_i}$, and the estimator (5) is clearly able to detect non-random distributions with a high level of significance. The scatter in the numerical ensemble can be judged from the size of the error bars, for the sake of clarity they have been shown in each plot only for a single case. For the plot of Fig. 1 the point distribution is that obtained from the $\{x\}_{-1}$ set. In this case the computation of the Q_l can be greatly reduced by using a random sub-sample of the particle set. This has been done also for other sets $\{x\}_s$, according to required computational load needed to evaluate the coefficients Q_l.

The results obtained have been found to be quite robust to changes in the size of the random sub-sample, usually an accurate evaluation of Q_l can be obtained already with a $\simeq 20\%$ random sub-sample of the particle set. How the Q_l change when particle populations with different clustering properties are selected is shown in Fig. 2. For a fixed $a(t)$ and R_c the coefficients Q_l have been evaluated for particle subsets associated with different thresholds. For any l the coefficients Q_l get larger when higher thresholds ν are selected. The choice of higher $\nu's$ corresponds to a more clustered distribution, and this is clearly detected by the coefficients Q_l. It is worth stressing that the trend is obtained according to the adopted definition (4) for $Q_l(\boldsymbol{r}_i)$. The angular power spectrum coefficients C_l scale approximately as $\propto N_i$, thus the enhanced clustering of particle subsets associated with higher $\nu's$ is masked for the C_l by the drop in N_i, with respect lower thresholds. The black squares in Fig. 2 are the rescaled linear theory coefficients (12). In this case the comparison has been made for the population $\delta \geq -1$, in order to avoid errors introduced by a linear bias approximation for Eq. (12), if subsets $\{x\}_s$ corresponding to higher thresholds ν would have been selected.

There is a good agreement with the numerical experiment up to $l \simeq 5$; beyond this angular scale ($\simeq \pi/5$) non-

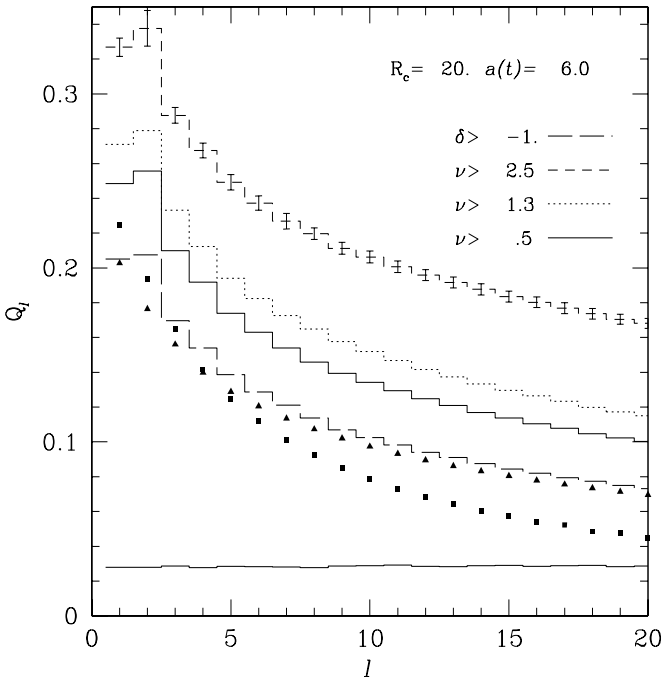

Fig. 2. Here the estimators Q_l have been computed at a fixed time and cut-off radius R_c from the point distributions generated by $\{x\}_s$, with $s = -1, 0.5, 1.3, 2.5$. The black squares are the rescaled linear theory coefficients C_l, computed from the CDM power spectrum at $a(t) = 6$ and for $R_c = 20$. The power spectrum has been normalized as in the simulations, so that the points compare directly with the $\delta = -1$ Q_l coefficients. The black triangles refer to the same theoretical coefficients, but computed in this case using the non-linear power spectrum of the simulation.

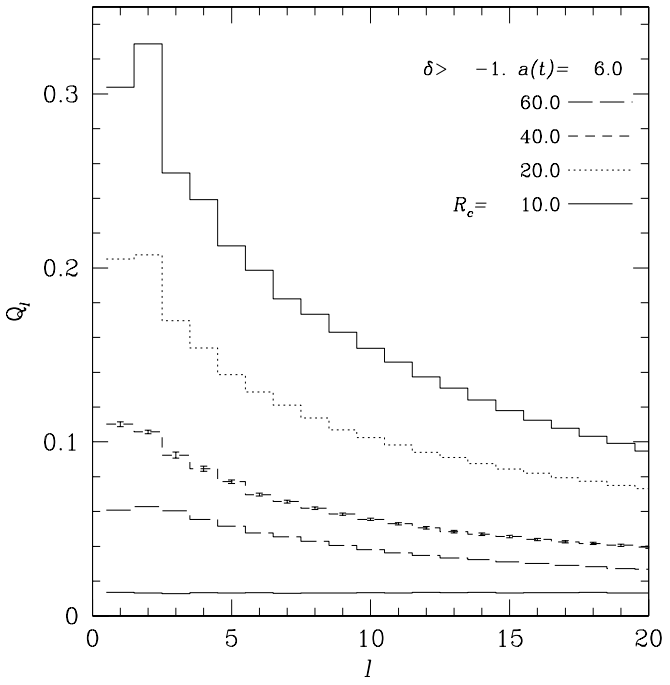

Fig. 3. The coefficients Q_l are shown at $a(t) = 6$ for four different cut-off radii $R_c = 10, 20, 40, 60$ (in $h^{-1}Mpc$ units). The estimators have been applied, as in Fig. 1, to the whole particle distribution.

linearity effects make the Q_l depart from the linearized coefficients. For the adopted normalization the size of the non-linear scale is about $R_{NL} \simeq 5 - 8h^{-1}Mpc$, at the epoch shown in the plot, which is in rough agreement with $R_c\Delta\theta$. When the power spectrum of the simulation particles is used to compute the theoretical coefficients there is an excellent agreement with the corresponding measured Q_l down to the smallest angular scale (black triangles of Fig. 2). For $l \leq 2$ the small differences show that the corrections (14) to Eq. (10), for the considered cosmological model are negligible already for $R_c \gtrsim 20h^{-1}Mpc$.

The coefficients Q_l are plotted for different cut-off radius R_c, in Fig. 3, for a fixed time and a given set $\{x\}_s$. There is a general tendency for the Q_l to get smaller at a fixed l when R_c is increased. This is expected because the bond distribution approaches isotropy as R_c gets higher. This trend is confirmed for all values of l and choices of different distributions $\{x\}_s$. Note also how the scattering in ensemble is reduced ($R_c = 40h^{-1}Mpc$) with respect to the smaller values of R_c shown in the previous plots. The analysis performed so far shows that the order parameters Q_l can be used as reliable statistical indicators of the global clustering properties of a point distribution.

From the histograms of Fig. 3 an interesting result is that the quadrupole coefficient Q_2 of the bond angular distribution is larger than the dipole coefficient Q_1. This is valid with a high significance only for values of the cut-off radii $R_c \lesssim 2r_0$, close the value of the 2-point clustering length. This excess of local anisotropies, compared to what is found for $R_c \gtrsim 20h^{-1}Mpc$, suggests that a significant amount of substructure is present in the particle distribution of dark matter halos. Jing (2000) has analyzed the halo dark matter density profiles from a set of cosmological $N-$body simulations. For those halos which are substructure rich and less virialized, he found large deviations from the analytical profile of Navarro, Frenk & White (1995). The above results therefore suggest that the proposed statistical method can be profitably used to correlate in a quantitative way the local halo anisotropies with other halo properties, when investigating their evolution in different cosmological models.

To analyze the clustering shape, more interesting results are obtained if the estimator (7) is applied to point distributions generated by different $\{x\}_s$. The total operation count of applying Eq. (7) to a given point distribution scales as $\propto N_p^4$, thus making it prohibitive to evaluate

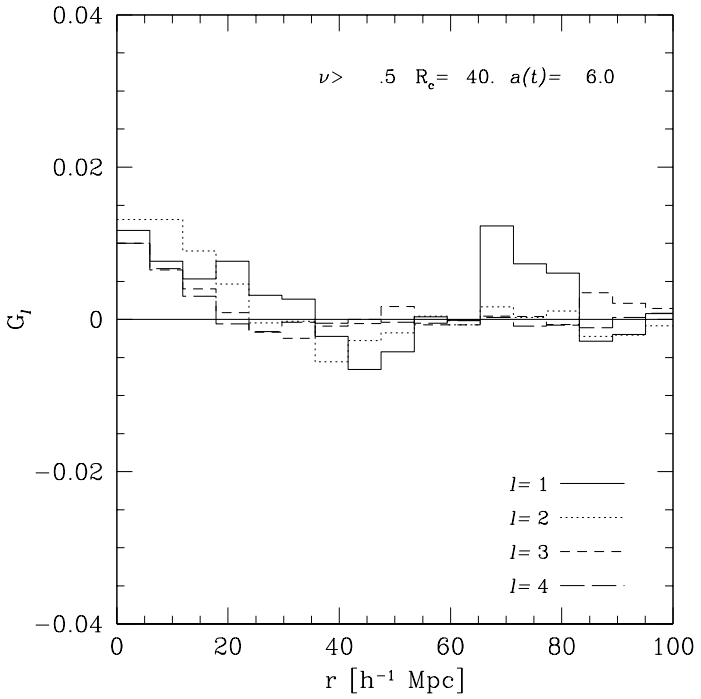

Fig. 4. The bond correlation functions $G_l(r)$, divided by G_0, are shown at $a(t) = 6$ and $R_c = 40$ for the distribution obtained from $\{x\}_{s=0.5}$. The different histograms are for different values of l, only G_l up to $l = 4$ are plotted. The line with a value for G_l close to zero is for a random distribution.

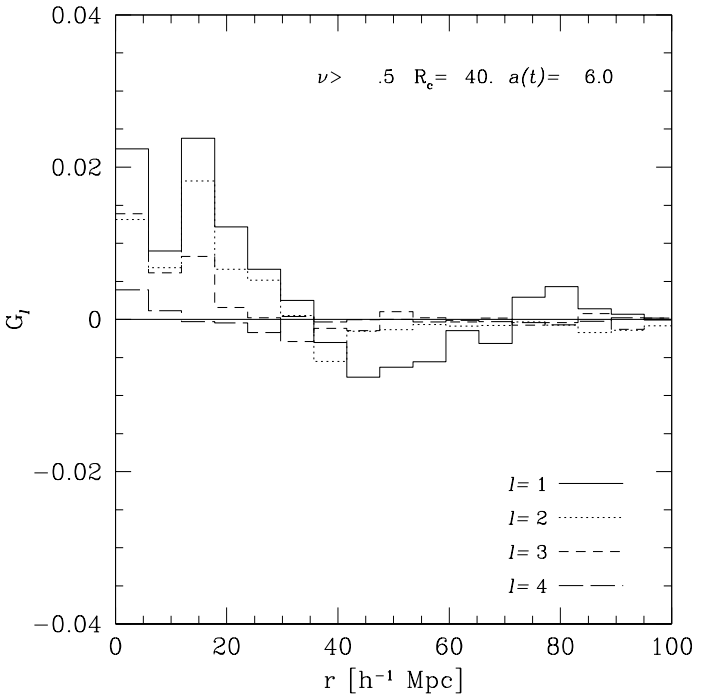

Fig. 5. As in Fig. 4, but in this case $G_l(r)$ have been evaluated summing in (7) over all the neighbors of the center particles; in the previous figure $G_l(r)$ were calculated taking random subsample of the neighbor sets.

$G_l(r)$ for the whole set $\{x\}_s$ of the simulation box. Therefore one has to resort to compute G_l taking as centers i in Eq. (7) a random subsample of $\{x\}_s$. The summations over the particles k at a distance r from i and the neighbor particles of the chosen centers ($\sum_j \sum_p$) are instead run over the whole set $\{x\}_s$, although a random dilution of the neighbor sets was performed in many cases in order to reduce the total computational cost. For a given particle i taken as a center the mean number $\bar{N}_i$ of neighbors j within a distance R_c from i ranges from $\simeq 10^2$ to $\simeq 10^4$, according to the chosen cut-off radius R_c, threshold level ν and expansion factor $a(t)$.

However the use of a random subset requires some care with respect to the previous procedures used to compute the coefficients Q_l. It has been found that the function $G_l(r)$ is subject, in some cases, to large sample-to-sample variations when different random subsets are chosen from a given $\{x\}_s$ to evaluate the corresponding G_l. This shows that $G_l(r)$ is a sensitive measure of the clustering patterns generated by a point distribution.

One of the consequences of this result is that the calculated $G_l(r)$ analyze in this case the clustering morphology of the point distribution as it is measured according to the selected random sub-sample, rather than that of the whole $\{x\}_s$. In order to consistently compare $G_l(r)$ for different choices of several parameters, the same random subset has been chosen at different times and for different R_c. Furthermore $G_l(r)$ for point distributions obtained for different thresholds have been calculated with random subsamples subject to the constraint that for a given $\{x\}_s$ the random subset was part of the random sample corresponding to a lower threshold, and so on down to the lowest level with $s = -1$. The application of these constraints leaves about $8,000$ particles in the simulation volume, which have been taken as center particles when the $G_l(r)$ were calculated. This procedure allows a consistent analysis to be made of the clustering morphologies detected by the $G_l(r)$ at different epochs and different values of the cut-off parameter R_c, angular scale l and threshold ν. The results obtained are shown in the following plots. Hereafter $G_l(r)$ is the estimator (7), applied to point distributions, divided by $G_0(r)$. Because of the large number of cases spanned in the parameter space, only a few of them have been plotted, the main features seen are quite general.

In Fig. 4 $G_l(r)$ is given as a function of r for different values of l, up to $l = 4$. The values for the other parameters are given in the panel. The correlation between angular distributions for different bonds decreases with increasing r and as smaller scales are probed, but with some exceptions. In this case G_l, for $l = 1$, has a significant

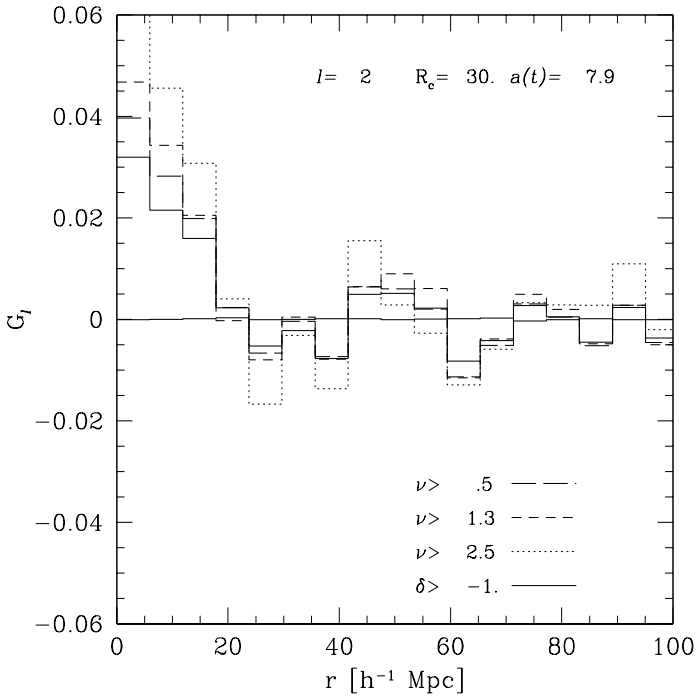

Fig. 6. The functions $G_l(r)$ plotted here have been obtained from different point distributions and the same values of l, R_c and $a(t)$ (given in the panel). The distributions are those given by the simulation particle subsets $\{x\}_s$, with $s = 0.5, 1.3, 2.5$, and $\{x\}_{s=-1}$.

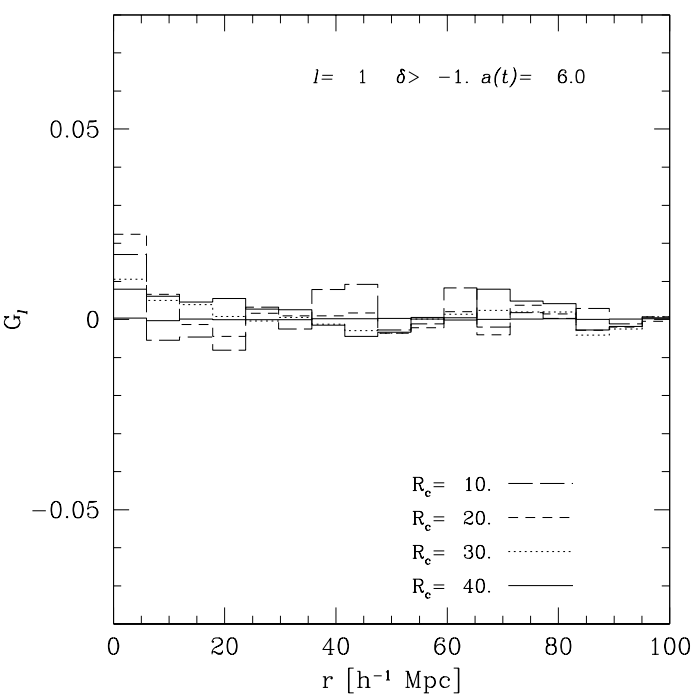

Fig. 7. The $G_l(r)$ are shown for different values of the cut-off radius: $R_c = 10, 20, 30, 40 h^{-1} Mpc$. The values of the other parameters are given in the panel.

correlation above the noise extending for a wide range of spatial distances. In other cases a similar trend has been found also for $l = 2$. $G_l(r)$ is a measure of the degree of coherence in the clustering patterns. For the example shown in the figure the result is a web of large-scale structures, in the examined point distribution, which is detected by the estimator. These features are detected at the lowest order only ($l \leq 2$), since their distribution in the bond spheres is limited to large angles. For $r \to 0$ the correlation strength increases because the two neighbor spheres overlap and partially trace the same structures. The continuum limit (21) is reached in the large sample regime $\bar{N}_i \to \infty$. For a finite number of points $G_l(r)$ measures the anisotropies of the distribution which samples the density field.

A theoretical estimate of how $G_l(r)$ are affected by noise terms is difficult, therefore an indirect check has been performed for the examples showed in the plots evaluating $G_l(r)$ in correspondence of a random distribution. For each of the considered cases this was obtained by considering tha same set of M_p pairs used in Eq. (7) to evaluate $G_l(r)$, but with a random distribution of $N_r = 10^3$ neighbors within each of the two spheres of radius R_c centered at the pair coordinates. The results show that $G_l \simeq 0$ for a random distribution of neighbors, so that the measured $G_l(r)$ are well above the noise.

For the example showed in Fig. 4 $G_l(r)$ have been computed using random sub-samples of the $\bar{N}_i \simeq 10^4$ neighbors of the center particles i. In this case the fraction f of selected neighbors was $f \approx 1/4$. Finite sample effects on the measured $G_l(r)$ can be estimated by repeating the evaluation of $G_l(r)$ for the case of Fig. 4, but this time running the summations in (7) over all the neighbors particles. The results are showed in Fig. 5, a comparison with the previous figure shows that the $G_l(r)$ profiles are qualitatively similar, suggesting that measurements of the clustering pattern by $G_l(r)$ are not affected by a dilution of the neighbor sets. Empirically it has been found that for the examples showed here this is valid for $f\bar{N}_i \gtrsim 10^2$.

$G_l(r)$ is also sensitive to different degrees of clustering, as shown in Fig. 6. In this case the G_l have been computed for different $\{x\}_s$ with the other parameters being kept constant (see the panel).

It is important to stress that this result was obtained because of the way in which the random subsets were chosen; if different subsets had been used for different $\{x\}_s$, then sample-to-sample variations would have dominated over the differences in G_l for different thresholds. A generic feature of G_l is a decrease in the signal as R_c is increased (Fig. 7). This is expected, since in going to higher cut-off radii, the point distribution in the bond sphere approaches isotropy.

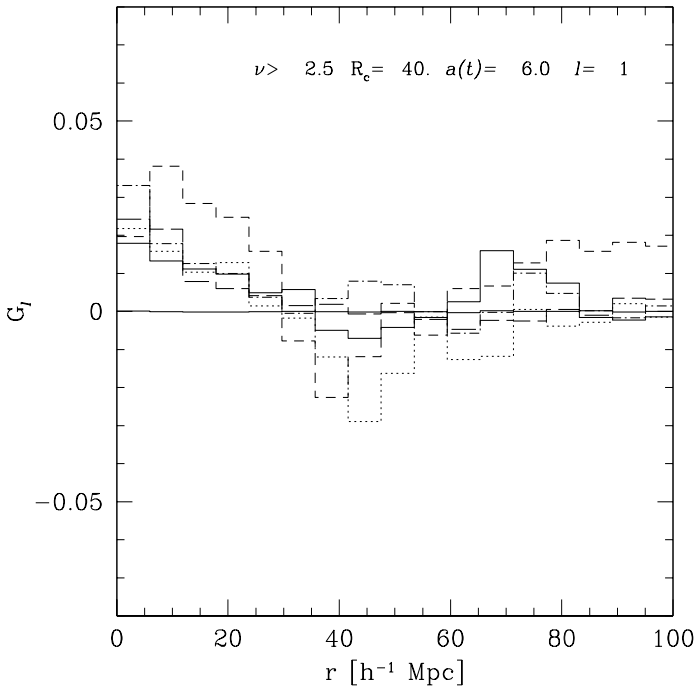

Fig. 8. The correlation functions $G_l(r)$ are plotted here for $l = 1, R_c = 40, a(t) = 6$ and point distributions obtained from $\{x\}_{s=2.5}$. Different histograms refer to $\{x\}_s$ from simulations with different random realizations of the ensemble.

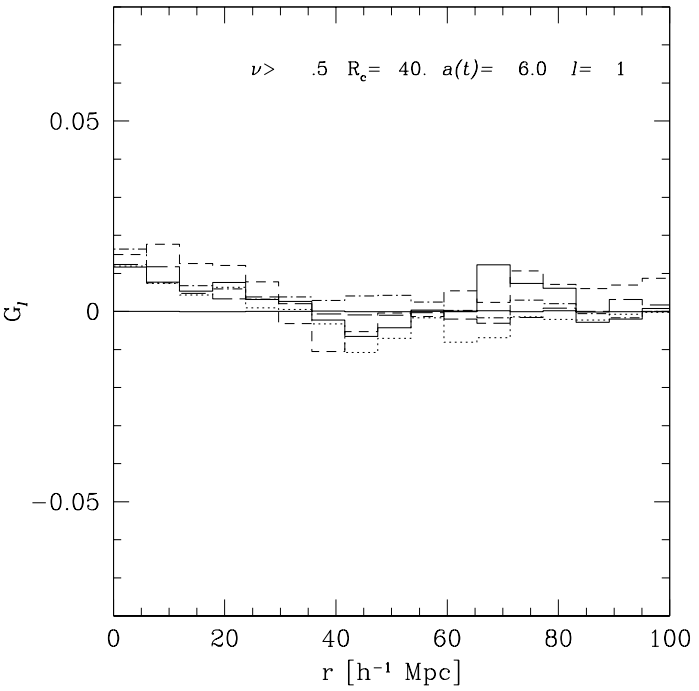

Fig. 9. The same as in Fig. 8, but for distributions obtained from $\{x\}_{s=0.5}$.

As previously stressed $G_l(r)$ has been found to be very sensitive to sample variations, the size of the fluctuations can be checked in Fig. 8, where $G_l(r)$ for a specified set of parameters has been computed for the different $\{x\}_s$ of the numerical ensemble. Because of the computational cost of evaluating $G_l(r)$ I have calculated these functions for the whole simulation ensemble only for $a(t) = 6$. The plot shows clearly how fluctuations in the ensemble, in this case, are of the same order as the variations in G_l shown in the previous plots when the input parameters were changed.

In one simulation large-scale correlations are clearly detected for $r \gtrsim 60h^{-1}Mpc$. These results are in accordance with those of Doroshkevich, Fong & Makarova (1998), who have analyzed the evolution of large-scale structure in standard CDM N-body simulations. Their analysis of sheets and filaments in the particle distribution is based upon the 'core sampling' method (Buryak, Doroshkevich & Fong 1994). A proper comparison is difficult because of the different statistical methods employed. The results obtained by Doroshkevich, Fong & Makarova (1998) reveal a richness of structures on scales of $\sim 10 - 100h^{-1}Mpc$, in agreement with the correlation range measured here by the functions G_l.

In Fig. 9 the functions $G_l(r)$ are shown, for different realizations, for the same example as in Fig. 8, but in this case for $\{x\}_{s=0.5}$ instead of $\{x\}_{s=2.5}$. The scatter in the ensemble is clearly reduced; this is a consequence of the lower threshold considered. For a point distribution approaching that of the whole mass in the simulation volume then $G_l(r)$ has a larger contribution from particles which are not part of filaments or of other large-scale structures. The signal is thus reduced as also are the variations between different random realizations. This is valid for cut-off radii much larger than the non-linear scale. If $R_c = 10h^{-1}Mpc$ then the functions G_l would have shown no sign of correlations, with strong fluctuations around zero independently of the separation distance r. The scattering in the ensemble is also reduced when smaller angular scales are probed (Fig. 10).

To summarize, the observed results have shown that the statistical method defined by the functions Q_l and $G_l(r)$ can be used to analyze the clustering morphology produced by gravitational clustering in a quantitative way. The method was originally devised in a completely different field for studying anisotropic structures and its application to cosmological clustering shows that, to some extent, it overlaps with previous analyses. The usefulness of the method has been tested by applying it to a set of cosmological N-body simulations with a CDM power spectrum. An application of the estimators to selected numerical outputs shows that the statistic is clearly able to discriminate between particle populations with different degrees of clustering. Several shape statistics have been in-

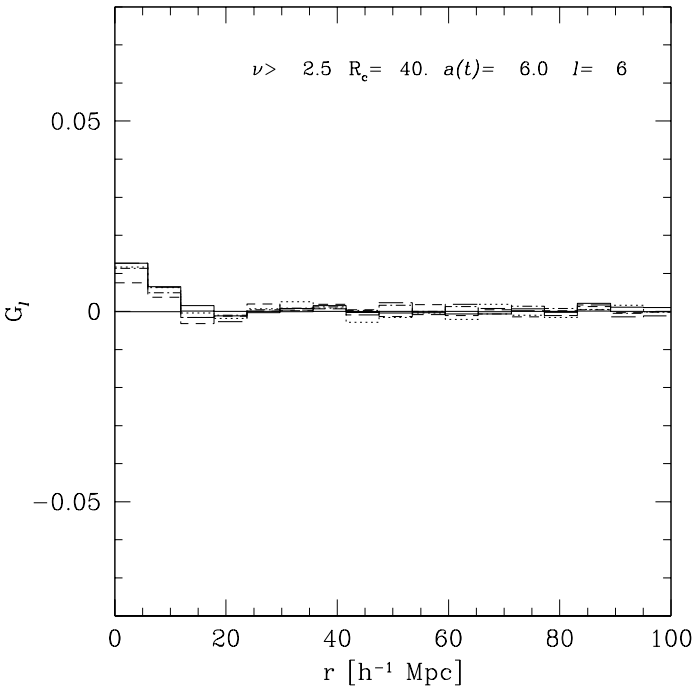

Fig. 10. As in Fig. 8, but for $l = 6$.

troduced to quantify the presence of filaments or sheet-like structures in the clustering network. The function $G_l(r)$ used here describes anisotropies in the clustering distribution by measuring the degree of correlation between the angular densities as seen from two different observers separated by r. $G_l(r)$ can then be considered a statistical measure of clustering patterns, with different scales probed by varying the input parameters l and R_c. This provides an alternative to other methods proposed so far to quantify the clustering morphology.

With large redshift surveys becoming available in the next few years, the proposed statistical method appears as a promising tool for analyzing patterns in the galaxy distribution. In practice real galaxy catalogs have an incomplete sky coverage which implies, for the observed harmonic coefficients, non-zero mean values and a coupling to different modes of the all-sky coefficients. This occurs because of the tensor window function of the angular mask arising from partial sky coverage. The required formalism has already been applied in the literature to angular catalogs (Scharf et al. 1992; Scharf et al. 1993) and its application to the Q_l coefficient is straightforward. Further complications arising from the radial selection function of the catalog and redshift space distortions have also been considered (Scharf et al. 1993; Ballinger, Heavens & Taylor 1995). For the functions $G_l(r)$ the analysis is much more cumbersome and is beyond the scope of this paper, which is intended to introduce the method and to illustrate its features by applications to numerical simulations.

An issue which has not been considered is whether this statistical tool can successfully be used to check the consistency of clustering data with cosmic models of structure formation. The analysis performed reveals (Fig. 8) that any method which quantifies in a sensible way the presence of filaments or planar structures in the cosmic web is also prone to cosmic variance. When different scales or levels of bias are considered, then the fluctuations are reduced (Figs. 9 & 10). The results obtained suggest that an application of the estimator $G_l(r)$ to real data will require different scales to be probed, in order to give useful constraints on cosmic theories.

Acknowledgements The author is grateful to the referee (M.Kerscher) for useful comments which improved the article.

References

Babul, A. & Starkman, G.D. 1992, ApJ 401, 28
Ballinger, W.E., Heavens , A.F. & Taylor, A.N. 1995, MNRAS 276, L59
Bardeen, J.M., Bond, J.R., Kaiser, N. & Szalay, A.S. 1986, ApJ 304, 15
Barrow, J.D., Bhavsar, S.P. & Sonoda, D.H. 1985, MNRAS 216, 17
Bartlett, J.G. astro-ph/9903260
Bertschinger E. & Gelb, J. 1991, Computers in Physics 5, 164
Bharadwaj, S., Sahni, V., Sathyaprakash, B.S., Shandarin, S.F. & Yess C. 2000, ApJ 528, 21
Bouchet, F.R., Strauss, M.A., Davis, M., Fisher, K.B., Yahil, A. & Huchra, J.P. 1993, ApJ 417, 36
Broadhurst, T.J., Ellis, R.S. , Koo, D.C. & Szalay, A.S. 1990, Nature 343, 726
Buryak, O.E., Doroshkevich, A.G. & Fong, R. 1994, ApJ 434, 24
Davé, R., *et al.* 1997, MNRAS 284, 607
da Costa , L.N., *et al.* 1994 , ApJ 424, L1
Doroshkevich, A.G., Fong, R. & Makarova, O. 1998, AA 329, 14
Efstathiou, G., Davis, M., Frenk C.S. & White, S.D.M 1985, ApJS 57, 241
Fisher, K.B., Davis, M., Strauss, M.A., Yahil, A. & Huchra, J.P. 1994, MNRAS 267, 927
Fry, J.N. 1984, ApJ 279, 499
Fry, J.N. 1986, ApJ 306, 366
Gangui, A. 1995, Ph.D. thesis, astro-ph/9512161, SISSA
Geller, M.J. & Huchra, J.P. 1989, Science 246, 897
Ghigna, S. *et al.*, 1997, ApJ 479, 580.
Goroff , M.H. et al. 1986, ApJ 311, 6
Gott, J.R., Dickinson, M. & Melott, A.L. 1986, ApJ 306, 341
Haile, J.H. & Gray, C.G. 1980, Chem. Physics Lett. 76, 583
Hu, W., Sugiyama N. & Silk, J. 1997, Nature 386, 37
Jenkins, A., *et al.* 1998, ApJ 499, 20
Jing, Y.P. 2000, ApJ 535, 30
Kerscher, M., *et al.* 1997, MNRAS 284, 73
Kerscher, M., *et al.* 1999, ApJ 513, 543
Klypin, A.A. & Shandarin, S.F. 1993, ApJ 413, 48
Landy, S.D., Shectman, S.A., Lin, H., Kirshner, R.D, Oemler, A.A. & Tucker, D. 1996, ApJ 456, L1
Luo, S. & Vishniac, E. 1995, ApJS 96, 429

Mc Donald, I.R. 1986, in *Molecular dynamics simulations of statistical mechanical systems*, Int. School of Physics E.Fermi, eds. G. Ciccotti and W.G. Hoover
Mecke, K.R., Buchert, T. & Wagner, H., 1994, AA 288, 697
Navarro J., Frenk C.S. & White, S.D.M. 1995, MNRAS, 275, 720
Park, C. 1991, MNRAS 251, 167
Peebles, P.J.E. & Hauser, M.C. 1974, ApJS 28, 19
Peebles, P.J.E. 1980, *The Large Scale Structure of the Universe,* Princeton University Press
Robinson, J. & Albrecht, A. 1996, MNRAS 283, 733
Ryden, B.S. *et al.* 1989, ApJ 340, 647
Saunders, W. *et al.* 1991a, Nature 338, 562
Saunders , W., *et al.* 1991b, Nature 349, 32
Scharf, C., Hoffman, Y., Lahav, O. & Lynden-Bell, D. 1992, MNRAS 256, 229
Scharf, C. & Lahav, O. 1993, MNRAS 264, 439
Schmalzing, J. & Buchert, T. 1997, ApJ 482, L1
Steinhardt, P.J., Nelson , D.P. & Ronchetti, M. 1983, PRB 28, 784
Szapudi, I. & Colombi, S. 1996, ApJ 470, 131
Totsuji, H. & Kihara, T. 1969, PASJ 21, 221
Ueda, H. & Yokoyama, J. 1996, MNRAS 280, 754
Ueda, H. & Itoh, M. 1999, ApJ 526, 560
Verde, L., Heavens, A.F. & Matarrese , S. 2000, MNRAS 318, 514
Vogeley, M.S., Geller, M.J., Park C. & Huchra, J.P. 1994, AJ 108, 745
Wang, Z.Q. & Stroud, D. 1991, J. Chem. Phys. 94, 3896
White, S.D.M. 1979, MNRAS 186, 145
White, S.D.M., Frenk, C.S., Efstathiou, G. & Davis, M. 1987, ApJ 313, 505
Zel'dovich, Ya. B. 1982, SvAL 8, 102